Building Data and Document-Driven Decision Support Systems

How do managers access and use large databases of historical and external facts?

A Data-Driven DSS

- Provides access to and manipulation of historical company data and in some cases external data.
- Simple file systems accessed by query and retrieval tools provide the most elementary level of functionality.
- Data warehouse systems provide additional functionality
- Data-Driven DSS with On-line Analytical Processing (OLAP) tools provide more functionality
- EIS and Spatial DSS

A Document-Driven DSS

- Document retrieval and analysis
- A collection of related, unstructured documents
- Search engines, document indexing and summarization
- Document and knowledge management

Comparing Data and Document-Driven DSS

- Structured vs. unstructured data
- Different analysis tools
- Storing data and documents
- Retrieval and indexing
- Fit with user needs
- What works when??

Four sub-categories of Data-Driven DSS

- Data Warehouses
- OLAP/Business Intelligence
- Executive Information Systemd
- Spatial DSS

Characteristics of Data Warehouses

- Subject-Oriented
  - Focuses on subjects related to business or organizational activity like customers, employers, and suppliers
- Integrated
  - Data is stored in a consistent format through use of naming conventions, domain constraints, physical attributes and measurements
- Time variant
- Non-volatile
  - Data does not change once it is in the warehouse

A Teradata Warehouse

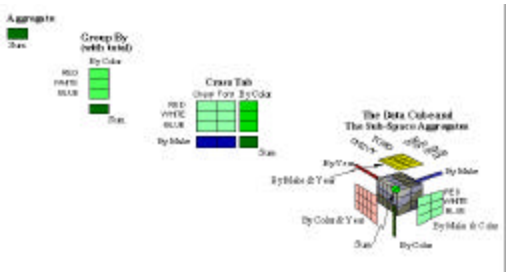
Teradata Data Warehouse Technical Overview

- Teradata RDBMS
  - designed for parallelism -- how many Aces?
  - supports complex decision support workloads
  - hardware and OS independent
  - the larger the data volume and the more complex the queries, the greater the need for parallelism
- NCR Worldmark Servers

On-Line Analytical Processing (OLAP)

- Can create various views and representations of data
- Fast Analysis of Shared Multi-dimensional Information (FASMI test)
- Multi-dimensional database
  - Captures and presents data as arrays that are "dimensioned" by relevant attributes

3-Dimensional Data Cube



Aggregating DSS Data

cross-tab

symmetric aggregation result

	Chevy	Ford	By Color
RED			
WHITE			
BLUE			
By Make			Sum

Executive Information Systems (EIS)

- Emphasis on graphical displays
- Present information from the corporate database
- Provide canned reports or briefing books to top-level executives
- EIS term has historical value in that such systems were developed separately from DSS and GDSS

Executive Information Systems

- EIS focuses on the decision-making information needs of senior managers
- Let's manager's "drill down" for more information and then try to manage data and avoid overload for the manager
- Enterprise-Wide DSS that helps senior managers analyze , compare, and highlight trends to recognize opportunities and/or problems

EIS vs. MIS

Spatial DSS

- A support system that represents large data sets using maps
  - Helps people access, display, and analyze data that have geographic content and meaning
  - Ex: Crime analysis and mapping, customer demographic analyses and political voting patterns by ward or precinct

Data-driven DSS vs. Business Intelligence (BI)

- BI is umbrella term used for software and systems to improve business decision making by using Data-Driven Decision Support Systems

Managers and Data-Driven DSS

- Managers want to find their own answers to business questions
- Managers are **NOT** willing to wait while financial or marketing analysts create special reports from databases
- Managers are the customers and advocates for data-driven DSS

Comparing DSS Data and Operating Data

- DSS Data - is data **ABOUT** transactions and occurrences
- Operating Data - is a record of specific business transactions
- These two types of data differ in five ways: Data Structures, Time Span, Summarization, Data Volatility, Data Dimensionality

DSS and Operating Data Differences

- Data Structures
  - DSS Data
    - Tables will need to be joined to complete a query
    - Do not include details of each transaction
    - Includes summaries of the transactions
    - May have data redundancies in the data structures if that will speed up queries - normalization is not required

Data Structures

- Operating Data
  - Both software and the hardware are optimized to support transactions about daily operations
  - Stored in many different tables
  - Stored data represents information about the specific transactions

More Differences

- Time Span
  - DSS Data
    - Is a snapshot of the operating data at given points in time
    - Historic time series of operating data
    - Storing multiple “time slices” of operating data
  - Operating Data
    - Is current and it shows the current status of business transactions

More Differences

- Summarization
  - DSS Data
    - Summarized in the database
    - Sometimes consist of exclusively derived data
  - Operating Data
    - DETAILED, It is not summarized in the database

Data Differences

- Data Volatility
  - DSS Data
    - Is non-volatile
  - Operational Data
    - Is volatile, data changes depending upon new transactions that occur

Data Differences

- Data Dimensionality
  - DSS Data
    - Has multiple dimensions
    - Always related from a manager’s and an analyst’s point of view
  - Operational Data
    - Has single dimension

Metadata

- "Data about data"
- Provides a directory to help decision support system DBMS locate contents of the data warehouse or data store
- Guide to mapping data as it is transformed from the operational environment to the data warehouse environment

Metadata

- Serves as a guide to the algorithms used for summarization of current detailed data
- Semantic information associated with a given variable
- Must include business definitions of the data and clear, accurate descriptions of data types, potential values, the original source system, data formats, and other characteristics

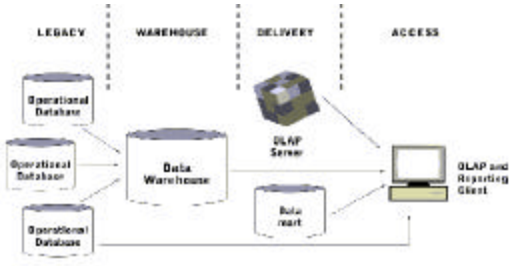
Summary of differences

Factors	Operating Data	DSS Data
Data Structures	normalized	integrated
Time Span	current	historical
Summarization	none	extensive in some systems
Data Volatility	volatile	non-volatile
Data Dimensions	one dimension	multiple dimensions
Metadata	desirable	required and important

An Interconnected Data-Driven DSS Architecture

- Data Store
  - Consists of one or more databases
- Data extraction and filtering component
  - ETL used to extract and validate the data
- Query and Reporting
- User Interface tool for data analysis and presentation

Data-Driven DSS Architecture



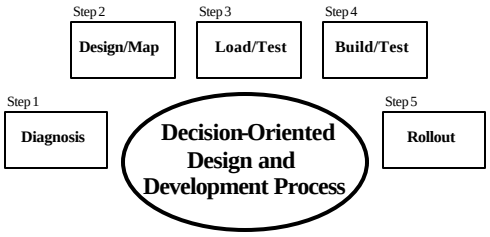
Implementing Data and Document-Driven DSS

- Issues that must be confronted
  - A DSS data store is not a static database - we add new time slices
  - Decision support infrastructure includes hardware, software, people, and procedures
  - Structure of the Data-Driven DSS data store and implementation must be examined in the context of the entire DSS infrastructure

Implementing Data-Driven DSS

- Technical Aspects
  - Must provide required analysis capabilities with acceptable query performance
  - Data-Driven DSS must support the data analysis needs of the decision-makers
  - Web-based, real-time systems

Data-Driven DSS Development Process



Developing an EIS

- Determine executive information requirements
- Staff group creates screens and information displays
- Group experiments with how data is presented and receive feedback from users
- User Interface issues are the “Key”

Developing Document-Driven DSS

- Initial Data Gathering or Diagnosis
- How will documents be stored?
- Search, retrieval, summarization and display
- How do you add value? Is retrieval and display enough?
- Tracking a historical record, finding patterns, making sense of the documents

Finding Success

- Identify an influential project champion
- Be prepared for technology shortfalls
- Tell everyone as much as you can about the costs
- Invest in training
- Market and Promote the Data or Document-Driven DSS

In Conclusion

- Different technologies are needed for systems built using structured data versus documents.
- Data-driven DSS have evolved from simple verification of facts to analysis of the data and now to sophisticated analysis of very large historical data sets.
- We are still innovating with Document-driven DSS.

Questions

- What is an EIS?
- When does a company need a data warehouse and OLAP?
- What is the FASMI test?
- How can managers increase the chances for success of a data-driven DSS project?